High Diversity Scheme for Wireless Networks based on Interference Cancellation

Chaitanya Rao and Babak Hassibi
Department of Electrical Engineering
California Institute of Technology
Pasadena CA 91125 USA
Email: {rao, hassibi}@systems.caltech.edu

***Abstract*— Consider a wireless network with m transmitter-receiver pairs and an additional n relay nodes to assist communication. We are interested in the rate/diversity trade-off of such a system. Since the presence of interference is known to reduce diversity significantly, we propose a transmission scheme based on interference cancellation by the relay nodes. This scheme achieves a diversity linear in the number of relay nodes (over all rates up to the maximum possible). Compared to a protocol where receivers decode all transmitted messages, the new scheme is seen to achieve higher diversity at higher rates.**

I. Introduction

In wireless networks reliable communication is limited by the presence of interference, fading and noise. In the high SNR regime we can measure reliability by diversity, the exponent of the error probability due primarily to fading. Relay nodes are known to provide diversity by setting up multiple independent transmission paths between each transmitter and receiver [1], [2]. In this way errors will only occur in the rare event that all of the links are in a deep fade.

In this work we look at a network with m transmitters each having a receiver they wish to communicate with. Assisting us in this communication are n relay nodes which provide redundant paths from sender to receiver. For reliable communication each receiver needs to decode its intended signal with low probability of error.

We will assume that the fading coefficients between any two communication nodes are independently and identically distributed according to a Rayleigh distribution. This may be a reasonable assumption in networks with high scattering and no direct line-of-sight transmission path, such as an indoor setting. The capacity of such channels (interference channels with relays) is an open problem, here we consider a particular transmission strategy and aim to achieve high diversity over a range of transmission rates.

We shall adopt the definition of rate and diversity in [3]: we consider a family of codes of fixed block length and increasing SNR and say that user i supports a multiplexing gain of r_i if its data rate $R_i(SNR)$ satisfies $\lim_{\mathrm{SNR}\to\infty} \frac{R_i(\mathrm{SNR})}{\log \mathrm{SNR}} = r_i$. Each user has a diversity d if the average error probability $\mathrm{P_e}$ for the system (all users) behaves as $\lim_{\mathrm{SNR}\to\infty} \frac{\mathrm{P_e}(\mathrm{SNR})}{\log \mathrm{SNR}} = -d$. Hence the higher value of d for a given multiplexing gain, the more reliable the system at that corresponding rate.

We will be defining a wireless scheme in which the diversity increases linearly in the number of relay nodes. The model has also been investigated in [4] with power efficiency in mind. We then find its diversity through an outage probability calculation. This will be expressed as a function of the sum of the multiplexing gains of the users. We postpone proof outlines to the final section.

The diversity-multiplexing gain trade-off was first defined and calculated for MIMO systems in [3]. In recent years cooperative wireless schemes have been analyzed [5], [6] and trade-off curves determined for a variety of networks incorporating relays [7], though the optimality of these curves over particular ranges of multiplexing gains remains an open problem. In this work we also compare the diversity-multiplexing trade-off curve to another scheme we considered in [8], in which there was no interference cancellation and receivers decoded all transmitted signals.

II. Model and Transmission Scheme

We have m transmitter-receiver pairs and n relay nodes, shown in Figure 1. On these nodes we shall impose the half-duplex condition [7]: they cannot simultaneously receive and transmit information. Let f_{ik}, g_{ij}, h_{jk} be fading coefficients between nodes, where $i = 1$ to m indexes the transmitters, $j = 1$ to n indexes the relays and $k = 1$ to m indexes the receivers respectively. All coefficients are assumed to be independent and identically drawn from a $\mathbb{C}\mathcal{N}(0, 1)$ complex Gaussian distribution. We assume each relay node has knowledge of all fading coefficients. This is a rather strong assumption. To justify it, suppose that none of the channels are changing rapidly and some form of

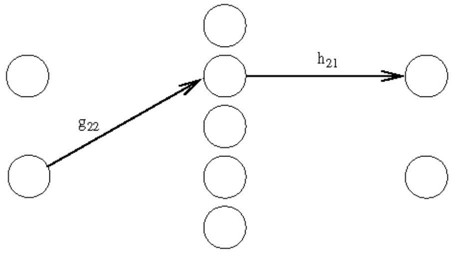

Fig. 1. Wireless network with m transmitter-receiver pairs and n relay nodes. In Stage 1 the transmitters send to the relays and receivers, in Stage 2 the relays forward a scaled version of what they have received to the receivers. Two of the fading coefficients are shown.

feedback is used. We also note that our results can be viewed as outer bounds for schemes that do not assume this channel knowledge. The receivers, however, need not have knowledge of all the channels.

Initially the m transmitter nodes send their signals simultaneously. Each of the relays and receivers obtains a faded linear combination of transmitted signals. The jth relay node then multiplies its received signal by a scalar d_j (a function of the fading coefficients) and forwards its resulting signal to the receivers while the m transmitting nodes are silent. Thus we have a two-stage process.

In Stage 1, if y_k denotes the signal received at the kth receiver and s_j the signal received at the jth relay node we may write

$$\begin{aligned} y_k^{(1)} &= \sqrt{P}\sum_{i=1}^{m} f_{ik}x_i + w_k^{(1)} \quad k = 1, 2, \ldots, m, \\ s_j &= \sqrt{P}\sum_{i=1}^{m} g_{ij}x_i + v_j \quad j = 1, 2, \ldots, n, \end{aligned}$$

where w_k and v_j are additive complex Gaussian noise at the kth receiver and jth relay node respectively. The constant P represents the power of each transmitted signal so as to normalize x_i: we assume $\mathbf{E}|x_i|^2 = 1$ and that each transmitter has the same average power P. Also assume w_k is zero-mean complex Gaussian noise with $\mathbf{E}|w_k|^2 = 1$.

In Stage 2, the kth receiver obtains a faded linear combination of signals from the relays:

$$\begin{aligned} y_k^{(2)} &= \sum_{j=1}^{n} s_j d_j h_{jk} + w_k^{(2)} \\ &= \sqrt{P}\sum_{j=1}^{n}\sum_{i=1}^{m} x_i g_{ij} d_j h_{jk} + \sum_{j=1}^{n} v_j d_j h_{jk} + w_k^{(2)}, \end{aligned}$$

for $k = 1, 2, \ldots, m$. In matrix form,

$$\begin{aligned} [(\mathbf{y}^{(1)})^T (\mathbf{y}^{(2)})^T] &= \sqrt{P}\mathbf{x}^T[F|GDH] + \mathbf{v}^T[0|DH] \\ &\quad +[(\mathbf{w}^{(1)})^T(\mathbf{w}^{(2)})^T], \end{aligned} \tag{1}$$

where f_{ik} is the (i,k) entry of F, g_{ij} is the (i,j) entry of G, D is diagonal with d_k as its kth entry, and H has h_{jk} as its (j,k) entry.

Stage 1 of transmission can be shown to add one to the diversity established by Stage 2—effectively it adds one independent path f_{kk} from sender to receiver. From now we focus on the diversity arising from Stage 2 of transmission.

Note that the m receivers receive both their desired signals as well as (potentially) $m-1$ interference terms. In [8] it has been shown that decoding signals assuming the interference terms are noise results in a diversity of zero, for any transmission rate. Therefore [8] focused on multiple access channel (MAC) decoding (see e.g. [9]) to obtain diversity results. Here the relays assist in eliminating the interference. Thus we choose D to satisfy:

$$G_{m\times n}D_{n\times n}H_{n\times m} = cI_{m\times m} \tag{2}$$

for a scalar c, and where $I_{m\times m}$ represents the $m \times m$ identity matrix. As mentioned earlier the relay nodes are assumed to have knowledge of the channels and so are able to compute D. Of course, since we have m^2 equations in (2) and n unknowns (not including c), this is only possible if $n \geq m^2$. We will presently assume this and later comment on the case $n < m^2$.

Then we have

$$\mathbf{y}^{\mathrm{T}} = \sqrt{P}c\mathbf{x}^T + \mathbf{v}^T DH + \mathbf{w}. \tag{3}$$

That is, each receiver obtains only its intended signal (plus additive noise). To ensure unit amplification by each relay on average we additionally impose the condition $\sum_{j=1}^{n} {d_j}^2 = n$. From this constraint, we find c as follows: solve the equation $G\hat{D}H = I$ for $\hat{D}$, then let $D = \hat{D}/||\hat{D}||$ which has the required norm n. Then $cI = GDH = G\hat{D}H/||\hat{D}||$ from which $c = 1/||\hat{D}||$. When $n > m^2$ there are infinitely many solutions to $G\hat{D}H = I$, and we will later argue that the optimal choice is the minimum norm solution. We assume each receiver has knowledge of the scalar c.

We remark that the power efficiency of this interference-cancellation scheme was studied in Section V-G of [4]. A similar interference cancellation scheme is considered in [10] where capacity scaling laws are derived in the asymptotic case of a large number of relays. This requires less channel knowledge at the relays, but at least m^3 relay nodes are required. Unlike [10], in this work we are primarily interested in diversity.

III. Main Result—Outage Behavior

We now analyze the error probability behavior of the interference cancellation scheme defined in the previous section. From (3) and recalling that each receiver knows c we can find the mutual information for the ith transmitter-receiver pair and say that that pair is in outage if for a given rate R_i the instantaneous mutual information is below R_i:

$$R_i > \log\left(1 + \frac{Pc^2}{1 + \sigma_v^2 ||(DH)_i||^2}\right), \qquad (4)$$

for $i = 1, 2, \ldots, m$. Here σ_v^2 represents the noise variance at each relay and $(DH)_i$ is the ith column of DH.

Assume we use a coding scheme which is *universal*, meaning that at high SNR an arbitrarily low probability of error can be achieved when the channel is not in outage [11]. If no such coding scheme exists, then the analysis which follows will provide a *lower bound* on the error probability (upper bound on diversity).

Hence from now we assume the dominant error event for the interference cancellation network is outage—that is one of the users' data rates cannot be supported due to fading.

It can be shown that since $||(DH)_i||^2$ has an exponentially tailed distribution, inequality (4) may be approximated by $c^2 < \epsilon$, where $\epsilon = O(1/P)$. That is, the ith user is in outage if

$$c^2 < \frac{k(2^{R_i} - 1)}{P} \sim P^{-(1-r_i)}$$

where $2^{R_i} = P^{r_i}$ and k is a constant.

For $0 < r_i < 1$ the right side of this expression is small for large P so we let $\epsilon_i := kP^{-(1-r_i)}$ and the outage probability is then upper bounded by the probability that $c^2 = 1/||\hat{D}||^2$ is less than ϵ_i.

If $n > m^2$ we can minimize this probability over the infinitely many choices of D satisfying (2) by choosing D to have minimal norm. We can rewrite the system $GDH = cI$ in the form $Ad = b$, where

$$A = \begin{bmatrix} g_{11}h_{11} & g_{12}h_{21} & \cdots & g_{1n}h_{n1} \\ g_{11}h_{12} & g_{12}h_{22} & \cdots & g_{1n}h_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ g_{11}h_{1m} & g_{12}h_{2m} & \cdots & g_{1n}h_{nm} \\ g_{21}h_{11} & g_{22}h_{21} & \cdots & g_{2n}h_{n1} \\ g_{21}h_{12} & g_{22}h_{22} & \cdots & g_{2n}h_{n2} \\ \vdots & \vdots & \vdots & \vdots \\ g_{21}h_{1m} & g_{22}h_{2m} & \cdots & g_{2n}h_{nm} \\ \vdots & \vdots & \vdots & \vdots \\ g_{m1}h_{1m} & g_{m2}h_{2m} & \cdots & g_{mn}h_{nm} \end{bmatrix}, \qquad (5)$$

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix}, \quad b = \mathrm{vec}(I_{m\times m}).$$

(The vec operator stacks the columns of a matrix, forming a column vector.) The d which minimizes $||d||^2 = ||D||^2$ in this underdetermined system of equations is found by applying the pseudoinverse:

$$d = A^*(AA^*)^{-1}b.$$

Hence

$$\min ||D||^2 = d^*d = b^*(AA^*)^{-1}b.$$

Determining the outage probability behavior at high SNR then amounts to finding $\Pr\left(\frac{1}{b^*(AA^*)^{-1}b} < \epsilon\right)$. It turns out that this is of order ϵ^{n-m^2+1} which leads to the following main result whose proof is outlined in Section V.

Theorem 1: Consider the two-stage interference cancellation scheme described by (1), with m transmitter-receiver pairs and n relay nodes (where $n > m^2$). Choose D to satisfy (2) and have norm $||D||^2 = n$. Then if the ith transmitter has multiplexing gain r_i, the maximum diversity of the system is at least $d = (n - m^2 + 2)(1 - \max_i r_i)$.

IV. Discussion

Theorem 1 implies that provided relay nodes allow for independent fades to and from the transmitter/receiver pairs, diversity can be made to grow linearly in the number of relay nodes for rates up to the maximum possible sum rate $S = m/2$.

Of course, we need not insist that all m nodes are transmitting at once. In fact, we cannot do so when $n < m^2$. In this case, one can choose $K < m$ source destination pairs and perform the interference cancellation scheme using the remaining $2m + n - 2K$ nodes as relays (which requires $2m + n - 2K \geq K^2$ or $K \leq \sqrt{2m + n + 1} - 1$). If we perform this for all $\binom{M}{k}$ possible combinations of K transmit/receive pairs, then each pair will be communicating a fraction K/m of the time. Thus, assuming equal multiplexing gains $r_i = r$, the sum-rate now is $S = rK/2$. We therefore have the following result.

Theorem 2: Consider the two staged interference cancellation scheme just described. Then, if the multiplexing gains of all users are equal, $r_i = r$, then for a given sum-rate S, the diversity is at least

$$d(S) = \max_{K \leq \sqrt{2m+n+1}-1} (2m+n-2K-K^2+2)(1-2S/K). \qquad (6)$$

A. A MAC-based scheme

For the sake of comparison, we consider a protocol previously analyzed in [8], where D no longer is chosen to cancel interference but rather a MAC-based decoder [9] is used to deal with interference. In this scheme, initially K of the m transmitters transmit their signals simultaneously as before while the corresponding K receivers listen. In the next K stages the remaining $2(m-K)+n$ nodes transmit what they received to one of the receivers by choosing their phases in such away that the signals are received coherently for that receiver. This is repeated K times for each receiver. This $(K+1)$-stage process then is repeated for all $\binom{m}{K}$ combinations of k transmit/receive pairs. If we assume all the multiplexing gains are equal the sum rate of this scheme is $S=\frac{K}{K+1}mr_i$. In [8] it is shown that such a scheme achieves diversity equal to the number of independent paths established between transmitter and receiver: $2(m-K)+n+1$.

Maximizing this over choices of K we obtain the following trade-off curve:

$$d(S)=\max_{K\in\{1,\ldots,m\}}(2m-2K+n+1)(1-\frac{1+K}{K}S). \quad (7)$$

Functions (6) and (7) are plotted together with the interference cancellation scheme in Figure 2 for the cases $m=2$, $n=6$ and $m=3$, $n=10$. We see that both curves match at low rates, at intermediate rates the MAC-based scheme achieves higher diversity, but the interference cancellation scheme is able to achieve diversity at higher rates. In fact, the main motivation for considering the scheme of this paper is that the MAC scheme allows only sum-rates up to $\frac{m}{m+1}$, whereas interference cancellation allows for transmission up to a rate of $S=\frac{\sqrt{2m+n+1}-1}{2}$.

V. Proof of Main Result

In this section we outline the proof of Theorem 1. We wish to show that if A and b have the form given in (5) then $\Pr\left(\frac{1}{b^*(AA^*)^{-1}b}<\epsilon\right)\sim\epsilon^{n-m^2+1}$. The final result follows from noting that the direct transmission of Stage 1 adds one to the diversity.

Order the singular values of A as $0<\sigma_1\leq\sigma_2\leq\ldots\leq\sigma_{m^2}$. Then the eigenvalues of $(AA^*)^{-1}$ are $\frac{1}{\sigma_1^2}>\frac{1}{\sigma_2^2}>\ldots>\frac{1}{\sigma_{m^2}^2}$. This implies

$$\frac{1}{b^*(AA^*)^{-1}b}\geq\frac{\sigma_1^2}{||\mathbf{b}||^2}\ =\ \frac{\sigma_1^2}{m}$$
$$\Rightarrow\Pr\left(\frac{1}{b^*(AA^*)^{-1}b}<\epsilon\right)\ \leq\ \frac{1}{m}\Pr\left(\sigma_1^2<m\epsilon\right).$$

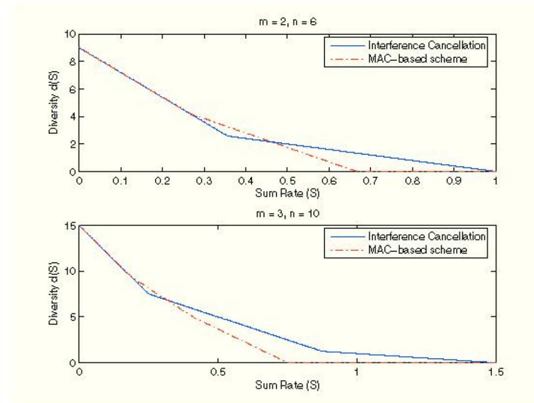

Fig. 2. Interference Cancellation and MAC-based Diversity-Rate Trade-off curves.

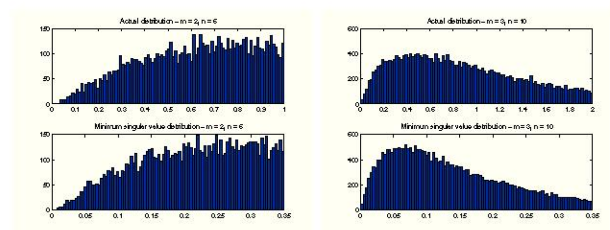

Fig. 3. Simulated probability density of $\frac{1}{b^*(AA^*)^{-1}b}$ (top plots) and the minimum singular value of A defined in (5) (lower plots). We claim that both have diversity $n-m^2+1$, which can be seen by the derivative at the origin behaving like ϵ^{n-m^2} ($m=2$, $n=6$ for the left plots, $m=3$, $n=10$ for the right plots).

By performing this bound we lose the structure of $\mathbf{b}$ but encouraged by Figure 3 the diversity is unchanged when making this approximation.

Now we wish to prove that $\Pr\left(\sigma_1^2(A)<\epsilon\right)\sim\epsilon^{n-m^2+1}$. The problem has reduced to identifying the distribution of the smallest singular value of A near the origin. If the matrix A were to have independent $\mathcal{CN}(0,1)$ entries, this result can be shown to be true by adopting the approach of [3] using the known joint distribution of the eigenvalues. In our case the entries of A in (5) are dependent and so computing the eigenvalue distribution is far more involved and we adopt an alternative approach.

The set of $m^2\times n$ matrices may be viewed as points in the vector space $\mathbb{C}^{m^2n}$. Matrices of the form A in (5) form a lower dimensional submanifold; denote this space by T. Assigned to T is a probability distribution induced by the complex Gaussian variables g_{ij} and h_{jk}. The matrices with smallest squared singular value less than ϵ will lie within a neighborhood of radius $\sqrt{\epsilon}$ of the submanifold of matrices having determinant zero (call this submanifold U and the neighborhood U_ϵ). Here distance is measured by the Frobenius (l^2) norm.

We are interested in the probability that a matrix from T is in this neighborhood of U. This will be given by integration of the probability density function of A over the region $T \cap U_\epsilon$.

For a given point $a \in T \cap U$ we identify the number of linearly independent directions from a normal to U and tangential to T. This requires knowledge of how the submanifolds T and U intersect, which in turn involves computation of their tangent spaces at a. This is based on the respective parametrizations of T and U at a.

All $m^2 \times n$ matrices of determinant 0 ($n > m^2$ so the rank is at most $m^2 - 1$) may be specified by $m^2 - 1$ of the vectors, with the remaining vectors being linear combinations of these. Hence we may write

$$A = [X|XG],$$

where X is an $m^2 \times (m^2 - 1)$ matrix and G is an $(m^2 - 1) \times (n - m^2 + 1)$ matrix specifying the linear combinations.

Observe that the columns of A are independent. For any $j \in \{1, 2, \ldots, n\}$ the $2m - 1$ values

$$g_{1j}h_{j1}, g_{1j}h_{j2}, \ldots, g_{1j}h_{jm}, g_{2j}h_{j1}, g_{3j}h_{j1}, \ldots, g_{mj}h_{j1}$$

specify any other entry of A. This is true by the identity

$$g_{ij}h_{jk} = \frac{g_{1j}h_{jk} \times g_{ij}h_{j1}}{g_{1j}h_{j1}}.$$

Hence by permuting rows as necessary a parametrization for A may be given by

$$A = \left[\frac{Y}{f(Y)}\right],$$

where Y is a $(2m-1) \times n$ matrix and $f(Y)$ is an $(m-1)^2 \times n$ matrix whose entries are of the form $y_{ij}y_{kj}/y_{1j}$ where i ranges from 2 to m, k ranges from $m+1$ to $2m-1$ and j ranges from 1 to n.

We carry out the integration over two stages. Firstly we integrate over $T \cap U_\epsilon$ these q dimensions. In the second stage we integrate over all $a \in T \cap U$. That is, the desired probability may be written in the form

$$\int_{a \in T \cap U} \int_{\substack{w \in T \\ ||w-a||^2 < \epsilon}} p(a, w)\sqrt{\det[J(a, w)J(a, w)^*]}\, dw da, \tag{8}$$

where $p(a, w)$ is the probability density function of Y evaluated at $(a, w) \in T$.

The determinant factor here represents the volume of an infinitesimal element in the m^2n-dimensional space: $J(a, w)$ is the matrix whose columns span the tangent space of $T \cap U_\epsilon$ at a. The inner integral is done in directions normal to U at a and may be approximated by that over a sphere of radius $\sqrt{\epsilon}$ centered at a. The pdf of matrices in T can be found explicitly (it depends on the modified Bessel function of the first kind). Furthermore the determinant in (8) is a ratio of two polynomials in its entries, so the asymptotic behavior of the integrand is dominated by the density function p. Knowledge of the asymptotic behavior of p enables us to upper bound the inner integral of (8) as $(\epsilon \ln(1/\epsilon))^{k(a)} g(a)$, where $g : T \cap U \to \mathbb{R}$ is some function independent of ϵ and $g(a) \sim \ln(1/||a||)$ for $||a|| \to 0$ and $g(a) \sim \exp(-||a||)$ as $||a|| \to \infty$. Furthermore one can show $k(a) = n - m^2 + 1$ almost surely.

Effectively the inner integral is over a k-dimensional small neighborhood of a and since the integrand is continuous, its value by the mean value theorem is close to the volume of that neighborhood times the value of the density function at a.

The outer integral becomes simple—the integrand is now bounded above by a constant multiple of $p(a)(\epsilon \ln(1/\epsilon))^{k(a)}$. Since $p(a)$ has the asymptotic properties of $g(a)$ and k is almost surely $n - m^2 + 1$, the outer integral will be some constant multiplied by $(\epsilon \ln(1/\epsilon))^{(n-m^2+1)} \sim \epsilon^{n-m^2+1}$ and we are done. That is, the maximum diversity of the system is $n - m^2 + 2$.

References

[1] A. Sendonaris, E. Erkip, and B. Aazhang. User cooperation diversity, Part I: System description. *IEEE Trans. Commun.*, 51:1927–1938, Nov. 2003.

[2] Y. Jing and B. Hassibi. Distributed space-time coding in wireless relay networks. *IEEE Trans. Wireless. Comm.*, 5:3524–3536, Dec. 2006.

[3] L. Zheng and D. Tse. Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels. *IEEE Trans. Inform. Theory*, 49:1073–1096, May 2003.

[4] A. F. Dana and B. Hassibi. On the power efficiency of sensory and *ad* hoc wireless networks. *IEEE Trans. Inform. Theory*, 52:2890–2914, Jul. 2006.

[5] N. Laneman, D. Tse, and G. Wornell. Cooperative diversity in wireless networks: Efficient protocols and outage behavior. *IEEE Trans. Inform. Theory*, 50:3062–3080, Dec. 2004.

[6] A. Høst-Madsen. Capacity bounds for cooperative diversity. *IEEE Trans. Inform. Theory*, 52:1522–1544, April 2006.

[7] K Azarian, H. El Gamal, and P. Schniter. On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels. *IEEE Trans. Inform. Theory*, 51:4152–4172, Dec. 2005.

[8] C. Rao and B. Hassibi. Rate-diversity trade-offs for interference channels with and without cooperation. *Proc. of Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, USA*, Oct. 2006.

[9] L. Zheng, P. Viswanath, and D. Tse. Diversity-multiplexing tradeoff in multiple-access channels. *IEEE Trans. Inform. Theory*, 50:1859–1874, Sep. 2004.

[10] H. Bölcskei, R. Nabar, O. Oyman, and A. Paulraj. Capacity scaling laws in MIMO relay networks. *IEEE Trans. on Wireless Commun.*, pages 1433–1444, Jun. 2006.

[11] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication.* Cambridge University Press, 2005.